



Covic, A., Keitel, C. , Porcu, E., Schröger, E. and Müller, M. M. (2017) Audio-visual synchrony and spatial attention enhance processing of dynamic visual stimulation independently and in parallel: A frequency-tagging study. *NeuroImage*, (doi:[10.1016/j.neuroimage.2017.08.022](https://doi.org/10.1016/j.neuroimage.2017.08.022))

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/145770/>

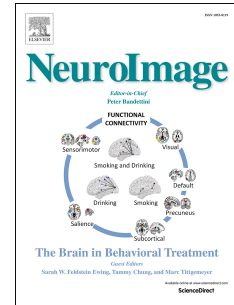
Deposited on: 11 August 2017

Enlighten – Research publications by members of the University of Glasgow  
<http://eprints.gla.ac.uk33640>

# Accepted Manuscript

Audio-visual synchrony and spatial attention enhance processing of dynamic visual stimulation independently and in parallel: A frequency-tagging study

Amra Covic, Christian Keitel, Emanuele Porcu, Erich Schröger, Matthias M. Müller



PII: S1053-8119(17)30669-9

DOI: [10.1016/j.neuroimage.2017.08.022](https://doi.org/10.1016/j.neuroimage.2017.08.022)

Reference: YNIMG 14253

To appear in: *NeuroImage*

Received Date: 20 April 2017

Revised Date: 1053-8119 1053-8119

Accepted Date: 6 August 2017

Please cite this article as: Covic, A., Keitel, C., Porcu, E., Schröger, E., Müller, M.M., Audio-visual synchrony and spatial attention enhance processing of dynamic visual stimulation independently and in parallel: A frequency-tagging study, *NeuroImage* (2017), doi: 10.1016/j.neuroimage.2017.08.022.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 TITLE

2 **Audio-visual synchrony and spatial attention enhance processing of dynamic visual**  
3 **stimulation independently and in parallel: a frequency-tagging study**

4

5 AUTHORS:

6 Amra Covic<sup>1,2\*</sup>, Christian Keitel<sup>3\*^</sup>, Emanuele Porcu<sup>4</sup>, Erich Schröger<sup>1</sup>, & Matthias M Müller<sup>1</sup>

7

8 AFFILIATIONS:

9 1 – Institut für Psychologie, Universität Leipzig, Neumarkt 9-19, 04109 Leipzig, Germany

10 2 – Institut für Medizinische Psychologie und Medizinische Soziologie, Universitätsmedizin  
11 Göttingen, Georg-August-Universität, 37973 Göttingen, Germany

12 3 – Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, University  
13 of Glasgow, 58 Hillhead Street, G12 8QB Glasgow, UK

14 4 – Institut für Psychologie, Otto-von-Guericke-Universität Magdeburg, Universitätsplatz 2  
15 Gebäude 23, 39106 Magdeburg

16

17 \* Joint first authors, equal contributions

18 ^ corresponding author: christian.keitel@glasgow.ac.uk

19

20 KEYWORDS:

21 spatial attention, selective attention, multisensory integration, audio-visual synchrony, brain  
22 oscillation, neural rhythm, steady-state response (SSR), EEG, brain-computer interface (BCI)

23

24 **ABSTRACT**

25 The neural processing of a visual stimulus can be facilitated by attending to its position or by  
26 a co-occurring auditory tone. Using frequency-tagging we investigated whether facilitation  
27 by spatial attention and audio-visual synchrony rely on similar neural processes. Participants  
28 attended to one of two flickering Gabor patches (14.17 and 17 Hz) located in opposite lower  
29 visual fields. Gabor patches further “pulsed” (i.e. showed smooth spatial frequency  
30 variations) at distinct rates (3.14 and 3.63 Hz). Frequency-modulating an auditory stimulus at  
31 the pulse-rate of one of the visual stimuli established audio-visual synchrony. Flicker and  
32 pulsed stimulation elicited stimulus-locked rhythmic electrophysiological brain responses  
33 that allowed tracking the neural processing of simultaneously presented stimuli. These  
34 steady-state responses (SSRs) were quantified in the spectral domain to examine visual  
35 stimulus processing under conditions of synchronous vs. asynchronous tone presentation  
36 and when respective stimulus positions were attended vs. unattended. Strikingly, unique  
37 patterns of effects on pulse- and flicker driven SSRs indicated that spatial attention and  
38 audiovisual synchrony facilitated early visual processing in parallel and via different cortical  
39 processes. We found attention effects to resemble the classical top-down gain effect  
40 facilitating both, flicker and pulse-driven SSRs. Audio-visual synchrony, in turn, only  
41 amplified synchrony-producing stimulus aspects (i.e. pulse-driven SSRs) possibly highlighting  
42 the role of temporally co-occurring sights and sounds in bottom-up multisensory integration.

43

44

45 **1. INTRODUCTION**

46 Behavioral goals, as well as the physical properties of sensory experiences, shape how neural  
47 processes organize the continuous and often rich influx of sensory information into  
48 meaningful units. One such process, selective attention, serves to prioritize currently  
49 behaviorally relevant sensory input while attenuating irrelevant aspects (Posner et al., 1980;  
50 Treisman and Gelade, 1980). In a visual search display, for example, items matching the  
51 color or orientation of a pre-defined target stimulus undergo prioritized processing relative  
52 to other items (Treisman and Gelade, 1980; Wolfe, 1994; Wolfe et al., 1989).

53 Another process exploits the spatial and temporal structure of dynamic sensory input,  
54 extracting regularities either in the visual modality alone (Alvarez and Oliva, 2009; Lee, 1999)  
55 or, by cross-referencing co-occurrences across sensory modalities (Fujisaki and Nishida,  
56 2005). In fact, aforementioned visual search can be drastically improved by presenting a  
57 spatially uninformative tone pip that coincides (repeatedly) with a sudden change in target  
58 appearance in a dynamic search array (Van der Burg et al., 2008).

59 This pop-out effect has been ascribed to a gain in relative salience of the target stimulus  
60 caused by the unique integration of auditory and visual information. The impression of a  
61 multisensory object hereby hinges on the temporal precision of coinciding unisensory inputs,  
62 also termed audio-visual synchrony, a critical cue for multisensory integration (Werner and  
63 Noppeney, 2011). Consecutive synchronous co-occurrences of the same auditory and visual  
64 stimulus components further increase the likelihood of multisensory integration (Parise,  
65 2012).

66 Generalizing this multisensory effect to our everyday experience of dynamic cluttered visual  
67 scenes, Talsma et al (2010) put forward that multisensory objects tend to involuntarily  
68 attract attention towards their position. As a consequence, they would gain an automatic  
69 processing advantage over unisensory stimuli. In a task that requires a sustained focus of  
70 attention on a specific position in the visual field multisensory stimuli may then act as strong

71 distractors (Krause et al., 2012) because they withdraw common processing resources from  
72 the task-relevant focus of attention.

73 Interestingly, this influence seems to work both ways: As Alsius et al. (2005) have shown  
74 focusing on a visual task impedes the integration of concurrent but irrelevant visual and  
75 auditory input. This effect has been related to the concept of the temporal binding window,  
76 a period during which co-occurring attended visual and auditory stimuli are most likely to be  
77 integrated (Colonius and Diederich, 2012). The window can expand for stimuli appearing at  
78 attended locations but remains unaffected (or contracts) when spatial attention is averted  
79 (Donohue et al., 2015).

80 Both phenomena - the involuntary orientation of spatial attention towards multisensory  
81 events as well as impeded multisensory integration when maintaining focused attention -  
82 have largely been studied in isolation (Talsma et al., 2010). We frequently encounter  
83 situations, however, in which the two biases can act concurrently. Moreover, they may  
84 fluctuate between having conjoined and conflicting effects depending on whether attended  
85 positions and multisensory events overlap or diverge in the visual field (that is in addition to  
86 their own inherent temporal variability (Keil et al., 2012).

87 This complex interplay therefore warranted a dedicated investigation in a paradigm that  
88 allowed contrasting both cases directly. In the present study, we manipulated trial by trial  
89 whether participants attended to a dynamic audio-visual synchronous stimulus while leaving  
90 a concurrently presented asynchronous stimulus unattended or vice versa.

91 We probed early cortical visual processing by tagging stimuli with distinct temporal  
92 frequencies (Norcia et al., 2015; Regan, 1989). This frequency-tagged stimulation elicited  
93 periodic brain responses, termed steady-state responses (SSRs). SSRs index continuous  
94 processing of individual stimuli in multi-element displays and have been demonstrated to  
95 indicate the allocation of spatial attention (Kim et al., 2007; Müller et al., 1998a; Walter et

96 al., 2012) as well as audio-visual synchrony (Jenkins et al., 2011; Keitel and Müller, 2015;  
97 Nozaradan et al., 2012).

98 Crucially, employing frequency-tagging allowed us to tease apart the relative facilitating  
99 effects of both factors as follows: Our paradigm featured two Gabor patches, one per lower  
100 visual hemifield, that each displayed two rhythmic physical modulations: As in classical  
101 frequency-tagging experiments they displayed a simple on-off flicker at different rates  
102 (14.17 and 17 Hz, respectively). Additionally, spatial frequencies of the Gabor patches  
103 modulated at slower rates (3.14 and 3.62 Hz, respectively), which gave the impression of a  
104 pulsation-like movement (see *Figure 1*). We exploited this pulsation to introduce audio-  
105 visual synchrony with a concurrently presented tone that carried a frequency modulation  
106 with the same temporal profile as one of the visual stimulus' movement (Giani et al., 2012;  
107 Hertz and Amedi, 2010 for similar approaches; see Keitel and Müller, 2015). Participants  
108 were then cued randomly on each trial to attend to one of the two stimulus positions, while  
109 one of the two Gabor patches pulsed in synchrony with the tone. This paradigm enabled  
110 comparisons of SSR-indexed visual processing between four cases of Gabor patch  
111 presentation: attended synchronous (A+S+), attended asynchronous (A+S-), unattended  
112 synchronous (A-S+) and unattended asynchronous (A-S-).

113 We expected our data to replicate well-described gain effects of top-down cued spatial  
114 attention on flicker-driven SSRs (Keitel et al., 2013; Kim et al., 2007; Müller et al., 1998a).  
115 Further, we assumed that these gain effects extend to pulsation-driven SSRs, because spatial  
116 attention should prioritize any information presented at an attended location.

117 Secondly, we hypothesized that in line with previous findings (Nozaradan et al., 2012) audio-  
118 visual synchrony produced gain effects on SSRs. In contrast to attentional gain, results of an  
119 earlier investigation suggested that synchrony-related gain effects may be specific to  
120 pulsation-driven SSRs. Using a paradigm similar to the present study, Keitel and Müller  
121 (2015) found that an SSR component with a frequency of twice the pulsation rate was

122 exclusively susceptible to synchrony-related gain effects. At this rate, the stimulation  
123 presumably contained strong transients critical for establishing audio-visual synchrony  
124 (Werner and Noppeney, 2011). If that were the case the current paradigm was expected to  
125 produce similarly selective effects. Alternatively, however, if audio-visual synchrony simply  
126 attracted spatial attention, then synchrony-related facilitation should mirror the pattern of  
127 attention-related gain effects on pulse- and flicker-driven SSRs. More specifically, synchrony  
128 alone should produce gain effects for flicker-driven SSRs.

129 Comparable patterns of attention- and synchrony-related facilitation would further point  
130 towards an account in which they may draw upon similar resources and therefore interact in  
131 facilitating visual processing: An attended stimulus would benefit less from audio-visual  
132 synchrony compared with an unattended synchronous stimulus, because attention has  
133 already been allocated to its position. Conversely, if attention- and synchrony-related  
134 facilitation relied on distinct neural resources, they were assumed to have independent  
135 additive effects on SSRs.

136 The latter finding could then be cast in a framework in which spatial attention biases are  
137 conveyed top-down via a fronto-parietal cortical network (Corbetta and Shulman, 2002),  
138 whereas audio-visual synchrony may have been established bottom-up via direct cortico-  
139 cortical connections or subcortical relays (Lakatos et al., 2009; van Atteveldt et al., 2014).

140 → *Insert Figure 1 here*

## 141 **2. METHODS**

### 142 **2.1. Participants**

143 We collected data from 14 participants with normal or corrected-to-normal vision and  
144 normal hearing. Participants gave informed written consent prior to experiments. None  
145 reported a history of neurological diseases or injury. They received course credit or a small  
146 monetary compensation for participation. The experiment was conducted in accordance



147 with the Declaration of Helsinki and the guidelines of the ethics committee of the University  
148 of Leipzig.

149 Two participants showed excessive eye movements during EEG recordings and were thus  
150 excluded. Data of 12 participants aged 18 – 31 years (all right-handed, 9 female) entered  
151 analyses. Previous studies have used comparable sample sizes to reliably (re)produce effects  
152 of spatial attention (Ding, 2005; Müller et al., 1998a; 1998b; Walter et al., 2015; Zhang et al.,  
153 2010) and audio-visual synchrony (Jenkins et al., 2011; Keitel and Müller, 2015; Nozaradan et  
154 al., 2012) on SSRs.

155

## 156 **2.2. Stimulation**

157 Stimuli were presented on a 19-inch cathode ray tube screen positioned 0.8 m in front of  
158 participants. The screen was set to a refresh rate of 85 frames per second and a resolution of  
159 1024 x 768 pixel (*width x height*). Visual experimental stimulation consisted of two  
160 monochrome Gabor patches with a diameter of  $\sim 3^\circ$  of visual angle, one located in the lower  
161 left and the other one located in the lower right visual field at eccentricities of  $4.5^\circ$  from  
162 vertical and  $2.5^\circ$  from horizontal meridians (see *Figure 1a*). Stimuli were presented against a  
163 grey background (RGB: 128,128,128; luminance =  $30 \text{ cd/m}^2$ ). Two black concentric circles ( $.4^\circ$   
164 of visual angle outer eccentricity, RGB: 0, 0, 0) in the center of the display served as fixation  
165 point.

166 Both Gabor stimuli underwent two independent periodic changes in the course of a trial:  
167 (1) The right patch presentation followed a cycle of 4 on-frames and 2 off-frames (2/1  
168 on/off-ratio) resulting in a 17 Hz flicker. The left patch flickered at a rate of 14.2 Hz achieved  
169 by repetitive cycles of 3 on-frames and 2 off-frames (3/2 on/off-ratio). (2) While flickering,  
170 the spatial frequency of the Gabor patches oscillated between a maximum of  $2 \text{ Hz}/^\circ$  and a  
171 minimum of  $1 \text{ Hz}/^\circ$  at a rate of 3.14 Hz for the right patch and 3.62 Hz for the left patch.  
172 Periodic spatial frequency changes gave the impression of alternating contractions and

173 relaxations that led to the percept of pulsing Gabor patches over time (*Figure 1c & d*). Pulse  
174 frequencies were chosen based on pilot experiments that served to determine a trade-off  
175 frequency range in which pulsing was readily perceptible, yet, still allowed driving periodic  
176 frequency-following brain responses (SSRs).  
177 In addition to the visual stimuli we presented a tone with a center frequency of 440 Hz  
178 binaurally via headphones. The frequency of the tone was rhythmically modulated following  
179 sinusoidal excursions from the center frequency (10% maximum excursion =  $\pm 44$  Hz). On  
180 each trial the modulation rate exactly matched the pulse rate of one of the two Gabor  
181 patches. Common rhythmic changes over time resulted in sustained audio-visual synchrony  
182 (see e.g. Schall et al., 2009).  
183 Prior to the experiment, we employed the method of limits (Leek, 2001) to approximate  
184 individual hearing thresholds using one of the experimental stimuli, a 3.14-Hz frequency  
185 modulated tone (see e.g. Herrmann et al., 2014; Keitel and Müller, 2015). In our  
186 implementation, participants listened to a series of 10 tone sequences with a maximum  
187 duration of 15 s per sequence. Tone intensity changed during each sequence while  
188 alternating between log-linear decreases and increases across sequences. Participants were  
189 instructed to indicate by button press when they stopped or started hearing respective  
190 tones. Cross-referencing button response times with tone intensity functions yielded  
191 individual estimates of psychophysical hearing thresholds, i.e. sensation levels (SL). In the  
192 experiment, acoustical stimulation was presented at an intensity of +35 dB SL.

193

### 194 **2.3. Procedure and Task**

195 Participants were seated comfortably in an acoustically dampened and electromagnetically  
196 shielded room and directed gaze towards the fixation ring on the computer screen. At the  
197 beginning of each trial, participants were cued to attend exclusively to the left or the right  
198 visual stimulus. To this end, a green semi-circle appeared inside the fixation ring for 500 ms

199 to indicate the task-relevant Gabor patch (see *Figure 1b*). Subsequently, the two pulsing  
200 Gabor patches and the tone were presented for 3500 ms. At the end of each trial, the  
201 fixation ring remained on screen for an extra 700 ms allowing participants to blink before  
202 the next trial started.

203 Participants were instructed to respond to occasionally occurring luminance changes of the  
204 cued Gabor patch (= targets) while ignoring similar events in the other patch (= distractors).  
205 During such events, Gabor patch luminance faded out to a minimum of 50% and back in  
206 within a 300 ms interval. Targets and distractors occurred in 50% of trials and up to 3 times  
207 in one trial with a minimum interval of 800 ms between subsequent onsets. Behavioral  
208 responses were recorded as space-bar presses on a standard keyboard. The responding  
209 hand was changed halfway through the experiment with the starting hand counterbalanced  
210 across participants.

211 We manipulated the two factors *attended position* (left vs. right Gabor patch) and audio-  
212 visual *synchrony* between attended Gabor patch and tone (synchronous vs. asynchronous) in  
213 a fully balanced design. Trials of the resulting four conditions – attended synchronous  
214 (A+S+), attended asynchronous (A+S-), unattended synchronous (A-S+) and unattended  
215 asynchronous (A-S-) – were presented in a pseudo-randomized order. Note that the tone  
216 was always in sync with one of the two Gabor patches. Therefore, in the two conditions in  
217 which the tone was out of sync with the attended Gabor patch, it was in sync with the  
218 unattended patch.

219 In total, we presented 600 trials (= 150 trials per condition) divided into 10 blocks (~5 min  
220 each). Before the experiment, participants performed training for at least one block. After  
221 each training and experimental block, they received feedback on the average hit rate and  
222 reaction time.

223

#### 224 2.4. Behavioral data recording and analyses

225 Responses were considered a 'hit' when the space bar was pressed between 200 to 1000 ms  
226 after target onset. We further defined false alarms as responses to distractors within the  
227 same time range. Based on these data, we calculated the response accuracy as the ratio of  
228 correct responses to the total number of targets and distractors for each condition and  
229 participant as follows:

$$230 \text{ ACC} = \frac{N_{\text{Hits}} + N_{\text{Correct Rejections}}}{N_{\text{Targets}} + N_{\text{Distractors}}} \quad [1]$$

231 where correct responses (= numerator) are the sum of target hits  $N_{\text{Hits}}$  and correctly rejected  
232 distractors  $N_{\text{Correct Rejections}}$ . Correct rejections were defined as the total number of presented  
233 distractors minus the number of false alarms. Accuracies were subjected to a two-way  
234 repeated measures analysis of variances (ANOVA) with factors of *attended position* (left vs.  
235 right Gabor patch) and *synchrony* (synchronous vs. asynchronous). Response speed,  
236 quantified as median reaction times, was analyzed accordingly.

237 For all repeated measures ANOVAs conducted in this study effect sizes are given as  $\eta^2$  (eta-  
238 squared). Where applicable, the Greenhouse–Geisser (GG) adjustment of degrees of  
239 freedom was applied to control for violations of sphericity (Greenhouse and Geisser, 1959).  
240 Original degrees of freedom, corrected  $p$ -values ( $p_{\text{GG}}$ ) and the correction coefficient epsilon  
241 ( $\epsilon_{\text{GG}}$ ) are reported.

242 Further Post-hoc tests – two-tailed  $t$ -tests for paired comparisons or against zero – were  
243 applied where necessary. We applied the Holm-Bonferroni procedure to correct  $p$ -values  
244 ( $p_{\text{HB}}$ ) for multiple comparisons (Holm, 1979).

245

#### 246 2.5. Electrophysiological data recording

247 EEG was recorded from 64 scalp electrodes that were mounted in an elastic cap using a  
248 BioSemi ActiveTwo system (BioSemi, Amsterdam, Netherlands) set to a sampling rate of 256  
249 Hz. Lateral eye movements were monitored with a bipolar outer canthus montage

250 (horizontal electrooculogram). Vertical eye movements and blinks were monitored with a  
251 bipolar montage positioned below and above the right eye (vertical electrooculogram). From  
252 continuous data, we extracted epochs of 3500 ms starting at audio-visual stimulus onset. In  
253 further preprocessing, we excluded 50% of epochs per condition (= 75) that corresponded to  
254 trials containing transient targets and distractors (= brief luminance fadings). These  
255 contained neural activity caused by processing target stimuli or motor activity due to  
256 response button presses that may have biased spectral estimates. Epochs with horizontal  
257 and vertical eye movements exceeding  $25 \mu\text{V}$  (=  $2.5^\circ$  of visual angle), or containing blinks  
258 were also discarded. To correct for additional artefacts, such as single noisy electrodes, we  
259 applied the 'fully automated statistical thresholding for EEG artefact rejection' (Nolan et al.,  
260 2010). This procedure corrected or removed epochs with residual artefacts based on  
261 statistical parameters of the data. Artefact correction employed a spherical-spline-based  
262 channel interpolation. For each participant FASTER interpolated up to 4 electrodes  
263 (median = 2) across recordings and an average of up to 5.6 electrodes (minimum = 1.9,  
264 median = 3.6) per epoch. Note that epochs with more than 12 artefact-contaminated  
265 electrodes were excluded from further analysis. In total, we discarded an average of 15% of  
266 epochs per participant and condition. Subsequently, data were re-referenced to average  
267 reference and averaged across epochs for each condition and participant, separately. Basic  
268 data processing steps such as extraction of epochs from continuous recordings and re-  
269 referencing made use of EEGLAB (Delorme and Makeig, 2004) in combination with custom  
270 routines written in MATLAB (The Mathworks, Natick, MA).

271

## 272 **2.6. Electrophysiological data analyses**

273 In our analyses we focused on two neural markers that have been repeatedly demonstrated  
274 to index attentional modulation: SSR amplitudes (Morgan et al., 1996; Müller and Hubner,  
275 2002; Quigley and Müller, 2014) and SSR inter-trial phase coherence (ITC, Kashiwase et al.,

276 2012; Kim et al., 2007; Porcu et al., 2013). Both measures also reflect effects of audio-visual  
277 synchrony on early visual processing (Nozaradan et al., 2012). Approaches to derive  
278 amplitudes and inter-trial phase coherence differ slightly and are thus described separately  
279 below. Both approaches required spectral decompositions of EEG time series for which we  
280 used the Fieldtrip toolbox (Oostenveld et al., 2011).

281

### 282 2.6.1. SSR power

283 Artefact-free epochs were truncated to segments of 3000 ms that started 500 ms after  
284 audio-visual stimulation onset and averaged separately for each EEG sensor, experimental  
285 condition and participant. The first 500 ms were omitted in order to exclude event-related  
286 potentials to stimulus onset from spectral analyses. From de-trended (i.e. linear trend  
287 removed) 3000 ms segments we quantified power (= squared amplitude) spectra by means  
288 of Fourier transforms. For the FFT, the 768 data points representing each 3000 ms segment  
289 were zero-padded to a length of 8192 ( $2^{13}$ ) to achieve a fine-grained spectral resolution  
290 (0.0312 Hz).

291 *Figure 2a* illustrates that our stimulation was effective in driving distinct SSRs: Power spectra  
292 pooled across all 64 scalp electrodes and experimental conditions showed clear peaks at the  
293 stimulation rates. Notably, spectra revealed strong harmonic responses at twice the pulse  
294 frequencies (6.28 and 7.24 Hz). We included these pulse-driven harmonics in further  
295 analyses because fundamental and harmonic responses have been hypothesized to reflect  
296 different aspects of stimulus processing (Kim et al., 2011; Pastor et al., 2007; Porcu et al.,  
297 2013) and showed modulation by synchrony in a previous study (Keitel and Müller, 2015).  
298 Grand-average topographical distribution of pulse-driven as well as flicker-driven SSR power  
299 averaged over conditions showed widespread maxima at parieto-occipital electrode sites  
300 (scalp maps in *Figure 2a*) that are typically observed in experiments with lateralized flicker  
301 stimulation (see e.g. Keitel et al., 2013).

302 For each participant and condition, SSR amplitudes were averaged across a cluster of 15  
 303 electrodes covering parieto-occipital maxima (Oz, O1, O2, I<sub>z</sub>, I<sub>1</sub>, I<sub>2</sub>, PO<sub>z</sub>, PO3, PO4, PO7, PO8,  
 304 P7, P8, P9, P10; as indicated in left-most scalp map in *Figure 2a*). Using a unified cluster of  
 305 electrodes across frequencies & stimuli allowed for a comparable spatial sampling of all SSR  
 306 components.

307 Amplitudes were further normalized by taking the decadic logarithm, then multiplying it by  
 308 20, to yield dB-scaled values (termed log-power in the following). All-positive SSR amplitude  
 309 values typically show a left-skewed distribution across participants. By taking their logarithm  
 310 we approximated a normal distribution (skew minimized) that better met the requirements  
 311 of parametric statistical procedures.

312 SSR log power was subjected to four-way repeated measures analysis of variances (ANOVAs)  
 313 with factors of driving *stimulus position* (left vs. right hemifield), *attention* (attended vs.  
 314 unattended), *synchrony* (synchronous vs. asynchronous) and *SSR component* (pulse 1f, pulse  
 315 2f and flicker 1f).

316 The factor *stimulus position* had no effect on SSR log power and did not show any interaction  
 317 with the other factors (see *Results*). This afforded collapsing normalized power across left  
 318 and right stimuli, i.e. across pulse frequency following ('pulse 1f') 3.14 Hz and 3.62 Hz, pulse  
 319 frequency doubling ('pulse 2f') 6.28 and 7.24 Hz, as well as flicker frequency following  
 320 ('flicker 1f') 14.17 and 17.00 Hz SSRs, respectively, in subsequent analyses.

321

### 322 2.6.2. SSR inter-trial phase coherence

323 We computed inter-trial phase coherence (Cohen, 2014) based on Fourier transforms of  
 324 artefact-free single trial epochs, truncated to 3000 ms segments (as described above for SSR  
 325 amplitude analyses) according to:

$$326 \quad ITC(f) = \left| \frac{1}{N} \sum_{n=1}^N \frac{c_n(f)}{|c_n(f)|} \right| \quad [2]$$

327 where  $c_n(f)$  is the complex Fourier coefficient of trial  $n$  at frequency  $f$  and  $|\cdot|$  indicates the  
328 absolute value. Inter-trial phase coherence as a measure of SSR modulation has been  
329 introduced to SSR analyses more recently (Kim et al., 2007; Nozaradan et al., 2012) and SSR  
330 amplitude and phase coherence have demonstrated different sensitivities to top-down  
331 influences on sensory processing (Kashiwase et al., 2012; Porcu et al., 2013). SSR Inter-trial  
332 phase coherence can be visualized as spectra that typically display narrow peaks at  
333 stimulation frequencies and higher order harmonics (Nozaradan et al., 2012; Ruhnau et al.,  
334 2016).

335 Similar to SSR amplitudes, ITCs showed broad topographic maxima at parieto-occipital  
336 electrode sites. Condition-averaged ITC spectra pooled across the 15-electrode cluster as  
337 described above (see section 2.6.1) revealed distinct peaks at the six frequencies of interest  
338 (Figure 2b).

339 Pooled ITCs were subjected to a four-way ANOVA with a design identical to SSR amplitude  
340 analyses. Note that ITCs were normalized by taking the natural logarithm prior to statistical  
341 evaluation. As for SSR log power, we found that ITC was insensitive to the *stimulus position*  
342 (left vs right; see section 3.2.2.), which again afforded collapsing across left- and right-  
343 stimulus driven in subsequent analyses.

#### 344 2.6.3. Power of the ongoing EEG and SSRs

345 As depicted in Figure 2c, SSRs have very low signal-to-noise ratios when being evaluated on  
346 the basis of averaged single-trial power spectra. Instead, these spectra accentuate the  
347 typical  $1/f^\alpha$  profile of power decreasing towards higher frequencies as well as peaks in the  
348 vicinity of 10 Hz that are consistent with alpha rhythmic brain activity. In turn, these features  
349 are much attenuated in SSR 'evoked' power and ITC spectra (Figures 2a and b).

350



## 351 2.6.4. Joint analyses of SSR amplitude and inter-trial phase coherence modulation

352 As laid out in the Results section, both of our manipulations, spatial attention and audio-  
 353 visual synchrony, revealed distinct patterns of effects on SSR amplitudes and ITCs. To further  
 354 characterize and compare these effects we computed an index that expressed attention-  
 355 and synchrony-related amplitude and ITC modulations for each subject and SSR frequency  
 356 component  $f$  (pulse 1f, pulse 2f and flicker 1f) according to:

$$357 \quad AMI_f = \frac{Amp_f^{att} - Amp_f^{ign}}{Amp_f^{att} + Amp_f^{ign}} \quad [3]$$

358 This attention modulation index (AMI) expressed the net gain effect of attention. AMIs were  
 359 calculated for each stimulus individually.  $Amp^{att}$  denotes SSR amplitudes when a stimulus  
 360 was attended and  $Amp^{ign}$  when the same stimulus was unattended (i.e. ignored). An  
 361 identically scaled synchrony modulation index (SMI) was computed by contrasting SSR  
 362 amplitudes between in-sync and out-of-sync conditions. We were thus able to compare both  
 363 indices directly. Entering ITCs instead of SSR amplitudes into formula (3) yielded ITC-based  
 364 AMIs and SMIs.

365 ANOVAs carried out for SSR amplitudes and ITC revealed that *attention* and *synchrony*  
 366 influenced SSRs additively, i.e. no interaction between these factors was found (see *Results*).  
 367 This finding justified collapsing AMIs across synchrony conditions and SMIs across attention  
 368 conditions for each SSR component, separately, in the following analyses. As an example, we  
 369 pooled the AMIs expressing the gain between synchronous conditions ( $A+S+$  vs  $A-S+$ ) and  
 370 asynchronous conditions ( $A+S-$  vs  $A-S-$ ).

371 Because further analyses rested firmly on the assumption of an absent *attention \* synchrony*  
 372 interaction, we additionally applied a Bayesian inference approach because in contrast to  
 373 the classical frequentist inference it allowed determining the amount of evidence in favor of  
 374 the null hypothesis ( $H_0$ : no interaction) explicitly. To this end, we estimated Bayes factors  
 375 (Rouder et al., 2012), i.e. the plausibility of a specific model given the data. First, separately

376 for SSR power and ITC, we determined models based on factors and interactions that turned  
377 out significant in ANOVAs. For example, SSR ITC was affected by a linear combination of  
378 factors *attention + synchrony + (synchrony \* SSR component)*. These models were tested  
379 against two alternative models, one including an interaction term (*attention \* synchrony*),  
380 and another one including a main effect of *stimulus position*.

381 The analysis was performed by means of the function *anovaBF* provided by the R (version  
382 3.3.0; R Core Team, 2013) package *Bayes factor* v0.9.12–2 (Morey et al., 2015). We adopted  
383 the Jeffrey-Zellner-Siow (JZS) prior with a standard scaling factor  $r$  of .707 (Rouder et al.,  
384 2012; 2009; Schönbrodt and Wagenmakers, 2015). Monte-Carlo resampling was based on  
385  $10^6$  iterations. Participants were considered as random factor. Importantly, Bayesian  
386 modelling favored the additive model (*attention + synchrony*) without an influence of the  
387 factor *stimulus position* (see *Results*) and further justified calculating AMIs and SMIs while  
388 collapsing across left and right stimuli. Results were robust against changing scaling factors.  
389 Finally, AMIs and SMIs were entered into a three-way ANOVA with factors of *SSR component*  
390 (pulse 1f, pulse 2f, and flicker 1f), *gain type* (attention vs synchrony) and *gain measure* (SSR  
391 amplitude vs ITC). Modulation indices were further tested against zero by means of t-tests  
392 (corrected for multiple comparisons).

393 → Insert Figure 2 here

### 394 3. RESULTS

#### 395 3.1. Behavioral data

396 Participants detected luminance fadings more accurately when attending to left Gabor  
397 patches (main effect *attended stimulus*:  $F(1,11) = 32.30$ ,  $P < 0.001$ ,  $\eta^2 = 0.579$ ; see *Table 1*).

398 Accuracy remained unaffected by in-sync vs. out-of-sync tone presentation (main effect  
399 *synchrony*:  $F(1,11) < 1$ ). The interaction of both factors was not significant ( $F(1,11) < 1$ ).

400 Reaction times increased slightly when participants performed the task on in-sync Gabor  
401 patches (main effect *synchrony*:  $F(1,11) = 9.27$ ,  $P < 0.05$ ,  $\eta^2 = 0.061$ ; see *Table 1*) but were

402 comparable between left and right stimuli (main effect *attended stimulus*:  $F(1,11) < 1$ ). As for  
403 accuracy, the interaction of both factors remained negligible ( $F(1,11) < 1$ ).

404 On average participants responded to 7.17% of distractors (median; interquartile range =  
405 14.00%). Due to their overall low occurrence false alarms were not analysed in detail. Note  
406 however that they contributed to the here employed accuracy score (see Formula 1).

407 → *Insert Table 1 here*

### 408 3.2. EEG data

409 We focused our analyses on SSR amplitudes and inter-trial phase coherence values (ITCs) to  
410 evaluate effects of spatial attention and audio-visual synchrony on early visual stimulus  
411 processing. Each stimulus drove three spectrally distinct SSR components: one at the  
412 frequency of stimulus pulsation, another one at twice the pulsation rate and a third  
413 following stimulus flicker (i.e., pulse 1f, pulse 2f and flicker frequencies, respectively).

414

#### 415 3.2.1. SSR power

416 SSR power decreased with increasing stimulus presentation rate (main effect *SSR*  
417 *component*:  $F(2,22) = 55.76$ ,  $P_{GG} < 0.001$ ,  $\epsilon_{GG} = 0.90$ ,  $\eta^2 = 0.301$ ; also see *Figure 3*) as has been  
418 documented extensively before (Keitel and Müller, 2015; Porcu et al., 2014). *Figure 3c*  
419 underlines that amplitudes further varied with the allocation of attention towards stimuli  
420 (main effect *attention*:  $F(1,11) = 24.15$ ,  $P < 0.001$ ,  $\eta^2 = 0.094$ ) and were affected by audio-  
421 visual *synchrony* ( $F(1,11) = 71.01$ ,  $P < 0.001$ ,  $\eta^2 = 0.067$ ). Amplitudes were comparable for  
422 left and right stimuli (main effect *stimulus position*:  $F(1,11) < 1$ ). A significant *SSR*  
423 *component \* synchrony* interaction ( $F(2,22) = 37.03$ ,  $P_{GG} < 0.001$ ,  $\epsilon_{GG} = 0.56$ ,  $\eta^2 = 0.057$ )  
424 warranted a closer investigation of synchrony effects on specific SSR components. The  
425 crucial *attention \* synchrony* interaction ( $F(1,11) = 1.12$ ,  $P = 0.313$ ,  $\eta^2 < 0.001$ ) as well as  
426 other interaction terms remained non-significant (maximum  $F(2,22) = 2.94$ ,  $P = 0.074$ ,  $\eta^2 =$   
427  $0.009$  for the *stimulus position \* SSR component interaction*).

428 The ANOVA results suggested a model based on the linear combination of factors *attention*  
 429 + *synchrony* + *SSR component* + (*synchrony* \* *SSR component*). Bayesian inference confirmed  
 430 that this model was more plausible than the model including an (*attention* \* *synchrony*)  
 431 interaction given our data ( $Bf_{\text{additive}} / Bf_{\text{interactive}} = 4.61 \pm 1.31\%$ ), as well as a model including a  
 432 main effect of *stimulus position* ( $Bf_{\text{additive}} / Bf_{\text{additive + stim. pos.}} = 7.55 \pm 2.47\%$ ).  
 433 The *SSR component* \* *synchrony* interaction originated from overall differences in the effect  
 434 of synchrony (in-sync minus out-of-sync) on each SSR component that was most pronounced  
 435 for pulse 2f components and virtually absent for flicker 1f responses (see *Figure 4a*). Specific  
 436 contrasts confirmed that pulse 2f SSRs were more susceptible to synchrony effects than  
 437 pulse 1f components ( $t(11) = 4.19, P_{\text{HB}} < 0.05$ ). Pulse 1f components in turn showed stronger  
 438 modulation than flicker 1f components ( $t(11) = 5.02, P_{\text{HB}} < 0.05$ ). Lastly, pulse 2f components  
 439 carried greater synchrony effects than flicker 1f components ( $t(11) = 7.83, P_{\text{HB}} < 0.05$ ).

440 → Insert Figure 3 here

### 441 3.2.2. SSR inter-trial phase coherence

442 ITC showed substantial variation with audio-visual *synchrony* ( $F(1,11) = 39.48, P < 0.001,$   
 443  $\eta^2 = 0.113$ ) and the allocation of *attention* ( $F(1,11) = 23.43, P < 0.001, \eta^2 = 0.139$ ) but no  
 444 effect of *SSR component* ( $F(2,22) = 2.24, P = 0.130, \eta^2 = 0.026$ ) or *stimulus position*  
 445 ( $F(1,11) < 1$ ). A significant *SSR component* \* *synchrony* interaction ( $F(2,22) = 16.16,$   
 446  $P_{\text{GG}} < 0.001, \epsilon_{\text{GG}} = 0.54, \eta^2 = 0.064$ ) indicated that some SSR components were more  
 447 susceptible to effects of audio-visual synchrony than others (*Figure 3b and d*). Remaining  
 448 interaction terms, especially the *attention* \* *synchrony* term ( $F(1,11) < 1$ ), failed to indicate  
 449 systematic effects (maximum  $F(1,11) = 2.80, P = 0.082, \eta^2 = 0.014$  for the *attention* \* *SSR*  
 450 *component interaction*). Only the *synchrony* \* *stimulus position* interaction was significant  
 451 ( $F(1,11) = 5.05, P = 0.046$ ) but explained a negligible amount of variance in the data  
 452 ( $\eta^2 = 0.003$ ) and was thus not further investigated. Note that the absence of effects of *SSR*  
 453 *component, stimulus position* or an interaction of both factors on ITC supports a comparable

454 spatial sampling (by averaging across a uniform cluster of 15 parieto-occipital electrodes; see  
455 *Methods*) of all SSR components.

456 Similar to SSR power, Bayesian inference supported the lack of an *attention \* synchrony*  
457 interaction. Comparing additive and interactive models by means of the Bayesian approach  
458 showed evidence in favor of the additive model ( $Bf_{\text{additive}} / Bf_{\text{interactive}} = 4.30 \pm 1.98\%$ ), again  
459 best modelled without an influence of the factor *stimulus position* ( $Bf_{\text{additive}} / Bf_{\text{additive + stim. pos.}}$   
460  $= 6.71 \pm 0.96\%$ ).

461 *Figure 4b* illustrates that the *SSR component \* synchrony* interaction stemmed from greater  
462 synchrony effects (in-sync minus out-of-sync) on pulse 1f than flicker 1f components  
463 ( $t(11) = 4.50$ ,  $p_{\text{HB}} < 0.05$ ). Also, synchrony affected pulse 2f ITC more strongly than flicker 1f  
464 components ( $t(11) = 5.06$ ,  $p_{\text{HB}} < 0.05$ ). Effects between pulse 1f and 2f SSRs were comparable  
465 ( $t(11) = 2.09$ ,  $p_{\text{HB}} = 0.19$ ).

466

### 467 3.2.3. Attention- vs Synchrony-related gain effects

468 As described in detail in the methods section, we computed indices that expressed SSR  
469 attention- and synchrony-related modulation of each SSR component. These modulation  
470 indices (AMIs and SMIs) allowed for a direct statistical comparison of the magnitude of  
471 attention and synchrony-related gain effects on SSR amplitudes and ITCs. As MI analyses  
472 assumed effects of attention and synchrony to be additive, further to the non-significant  
473 *attention \* synchrony* interactions reported above, we estimated the plausibility of additive  
474 vs interactive models given our data by using a Bayesian approach. The estimated Bayes  
475 factors for SSR power and ITC (see sections 3.2.1. and 3.2.2.) indicated that both results  
476 were more than 4 times more likely under the additive than the interactive model.  
477 Comparing modulation indices based on SSR amplitudes (*Figure 4E*) and SSR inter-trial  
478 coherence (*Figure 4F*) revealed that, overall, attention led to stronger gain effects on SSRs  
479 than synchrony ( $15.7\% \pm 1.8$  vs  $13.7\% \pm 1.8$ , mean  $\pm$  standard error; main effect *gain type*:

480  $F(1,11) = 28.79, P < 0.001, \eta^2 = 0.20$ ). Most importantly, however, this difference in gain  
481 effects varied between SSR components (interaction *gain type* \* *SSR component*:  
482  $F(2,22) = 6.66, P_{GG} = 0.007, \epsilon_{GG} = 0.898, \eta^2 = 0.13$ ) in the absence of a modulation of gain  
483 effects across *SSR components* alone (main effect:  $F(2,22) = 0.41, P = 0.668$ ).

484 From a methodological perspective it should be noted that power-based modulation  
485 indicated a small but significantly higher gain than ITC based modulation (main effect *gain*  
486 *measure*:  $F(1,11) = 19.77, P < 0.001, \eta^2 < 0.01$ ), an effect that further depended on whether  
487 attention or synchrony caused the modulation (interaction *gain measure* \* *gain type*:  
488  $F(1,11) = 7.85, P = 0.017, \eta^2 < 0.01$ ).

489 However, we disregarded these small effects to investigate the *gain type* \* *SSR component*  
490 interaction more closely. First, SSR amplitude and ITC-based modulation indices were tested  
491 against zero. Attention systematically modulated all SSR components (see Figures 4E & F;  
492 asterisks denote significant deviations from zero at a Holm-Bonferroni corrected alpha level  
493 of  $P < .05$ ). Synchrony, in turn, only modulated pulse 2f, but not pulse 1f and flicker 1f  
494 responses for both, SSR power- and ITC- based modulation indices.

495 Given these highly similar patterns we pooled across measures. Then we tested gain  
496 differences (Attention minus Synchrony) between SSR components. Elucidating the *gain*  
497 *type* \* *SSR component* interaction, gains differed more for flicker 1f than for pulse 1f SSRs  
498 ( $t(11) = 3.03, P_{HB} < .05$ ) and for pulse 2f SSRs ( $t(11) = 3.06, P_{HB} < .05$ ). In turn, gain differences  
499 were statistically comparable between pulse 1f and pulse 2f SSRs ( $t(11) = -0.92, P = .376$ )  
500 highlighting the exclusive role of the flicker-driven signal component.

501 → Insert Figure 4 here

#### 502 4. DISCUSSION

503 The role of top-down attention in multisensory binding and, conversely, bottom-up  
504 multisensory influences on attentional orienting have been studied largely independent of  
505 each other (Talsma et al., 2010). The present study was designed to bridge this gap.

506 Specifically, we studied situations in which participants attended to the position of one of  
507 two pulsing and flickering stimuli providing it with a top-down processing advantage over  
508 the other stimulus. Additionally, a tone pulsing in synchrony with either the attended or  
509 unattended stimulus was introduced to produce a strong multisensory bottom-up bias in  
510 visual processing. EEG-recorded SSRs driven by stimulus flicker and pulsation allowed us to  
511 test whether and how spatial attention and audio-visual synchrony acted, and possibly  
512 interacted, to facilitate cortical visual stimulus processing.

513

514 We evaluated two commonly used SSR measures, evoked power and inter-trial phase  
515 coherence (ITC) to quantify modulations in stimulus processing. Both measures widely agree  
516 on patterns of effects and will thus be considered jointly in the following.

517

518 Briefly summarizing the results, spatial attention facilitated pulse- *and* flicker-driven SSRs. In  
519 contrast, synchrony specifically facilitated pulse-driven SSRs only with greater effects on  
520 pulse 2f components while leaving flicker 1f components unaffected. Most importantly,  
521 attention and synchrony produced independent additive gain effects. We confirmed that,  
522 given our data, an additive model of both influences was more plausible than assuming  
523 interactive effects. These findings replicate results from an earlier study using a related  
524 paradigm. In that study we tested concurrent influences of feature-based attention and  
525 audio-visual synchrony on two spatially super-imposed Gabor patches (Keitel and Müller,  
526 2015).

527

#### 528 **4.1. Spatial attention facilitates processing of all stimulus aspects**

529 The described effects of spatial attention are in line with numerous studies demonstrating  
530 sensory gain effects on SSR-indexed cortical visual processing (Müller et al., 1998a; Störmer  
531 et al., 2014; Walter et al., 2015). Interestingly, our results show that spatial attention has

532 comparable effects on SSRs driven by two different but simultaneous rhythmic changes in  
533 stimulus appearance: a relatively fast on-off flicker ( $> 14$  Hz) and a slow-paced sinusoidal  
534 spatial frequency modulation (3 – 4 Hz). These results support the notion that spatial  
535 attention prioritizes all aspects of sensory information within its focus (Andersen et al.,  
536 2008; Keitel and Müller, 2015) as is central to psychological (Treisman and Gelade, 1980;  
537 Wolfe, 1994) and neurophysiological models of attention (Bundesen et al., 2015; Reynolds  
538 and Heeger, 2009) .

539 Note that participants performed better in the visual detection task when they attended to  
540 the left stimulus. This effect could be due to a left-hemifield advantage as has been  
541 described previously for rapid serial visual presentation paradigms (Śmigasiewicz et al.,  
542 2014; Verleger et al., 2011). In turn, SSR analyses did not show differences in stimulus  
543 processing between left and right stimulus positions. It is therefore possible that the  
544 imbalance in task performance did not stem from differences in early visual processing of  
545 left and right stimuli but was introduced at a later processing stage.

546

#### 547 **4.2. Synchrony selectively facilitates stimulus aspects relevant for multisensory integration**

548 Facilitation of visual processing by audio-visual synchrony has largely been studied using  
549 transient stimuli (Busse et al., 2005; Talsma et al., 2009). So far, only a few studies have  
550 demonstrated synchrony-driven effects while employing dynamic ongoing stimulation  
551 (Keitel and Müller, 2015; Nozaradan et al., 2012; Schall et al., 2009). Prolonged exposure to  
552 synchronous sensory input, however, can be a vital factor in multisensory integration  
553 because it improves the estimate of temporal correlations between visual and auditory  
554 stimuli over time (Parise and Ernst, 2016). This is important in situations with multiple  
555 concurrent stimuli (as studied here) because even unrelated visual and auditory events can  
556 occur simultaneously occasionally.

557



558 Our study corroborates this role of ongoing audio-visual synchrony. Interestingly, synchrony-  
559 related gain effects were thereby restricted to SSR components that reflected stimulus  
560 pulsing, i.e. those rhythmic modulations that produced the impression of synchrony.

561

562 Visual stimulus dynamics either matched with or differed from the spectral profile of the  
563 auditory stimulus, thus providing either maximal or minimal temporal correlation. Less  
564 intuitively, the SSR component at twice the pulsation rate (pulse 2f) showed greater  
565 synchrony modulations than the pulse-frequency following response (pulse 1f). In line with  
566 Keitel et al. (2015), who employed a stimulus with similar dynamic properties, the pulse 2f  
567 modulation was accounted for by the transients elicited by the stimulus at twice the  
568 stimulus pulsation rates during maximum up- and down-slopes of the sinusoidal modulation,  
569 or alternatively its extrema, i.e. peaks and troughs.

570

571 We propose that successive cross-modal phase resets may be the neural process underlying  
572 synchrony-related modulation of both pulse-driven components. Cross-modal phase  
573 resetting has been considered as the primary channel for multisensory interactions between  
574 early sensory cortices (Lakatos et al., 2009; van Atteveldt et al., 2014). Unlike neurons in  
575 higher order cortices, which are intrinsically multisensory (and hence sensitive to combined  
576 multisensory information) neurons in early sensory cortices are primarily sensory specific,  
577 but crucially sensitive to temporal information conveyed also by non-specific modalities. As  
578 underlined by Lakatos et al. (2008), appropriately timed inputs in one modality can aid in  
579 processing a stimulus presented in a different modality. In our case these connections may  
580 support phase stability of visual SSRs by providing a cross-modal temporal scaffold (Kayser et  
581 al., 2010; Lakatos et al., 2009). As a consequence, the temporal precision of cortical stimulus  
582 representations increases, which awards them a processing advantage (Chennu et al., 2009).

583

584 Although our results are broadly in line with Nozaradan et al. (2012), who firstly measured  
585 synchrony effects on SSRs, it is worth noting a discrepancy: In contrast to our findings the  
586 authors reported an effect on a flicker-driven SSR with a frequency of 10 Hz, while  
587 establishing synchrony with auditory beats at either 2.1 or 2.4 Hz. These differences may be  
588 accounted for by the fact that the authors presented only one visual stimulus centrally. In  
589 this setup, gain effects cannot not unambiguously be ascribed to synchrony, or alternatively,  
590 altered attentional demands between synchronous and asynchronous conditions.

591

#### 592 **4.3. Facilitatory effects of spatial attention and synchrony add up**

593 We found that attended and unattended stimulus experienced comparable gain through  
594 synchrony. Vice versa, synchronous and asynchronous stimuli were similarly facilitated when  
595 their position was attended. Remarkably, these findings point towards a dual reign of  
596 attention and audio-visual synchrony in early sensory cortices, suggesting that both  
597 influences can work independently and in parallel. This result seemingly contradicts previous  
598 studies (Alsius et al., 2005; Fairhall and Macaluso, 2009) that showed an interdependence  
599 between attention and multisensory interactions. However, this contradiction can be  
600 reconciled by examining the experimental paradigm employed in the current study.

601

602 Unlike previous experiments, in which mutual input from different senses was essential for  
603 successful behavioral performance, it is hard to construe a direct benefit from audio-visual  
604 synchrony in performing our task, i.e. the purely visual detection of luminance changes. Our  
605 paradigm might thus have promoted the independence between attention and audio-visual  
606 interactions triggering two concurrent, but distinct processes: On the one hand, performing  
607 the detection task required a sustained goal-driven deployment of spatial attention, while  
608 on the other hand merging the audio-visual signals was most likely a stimulus-driven

609 process, triggered by the high temporal correlation between auditory and visual signal  
610 components.

611

612 For these two processes to co-occur independently, we assumed the involvement of distinct  
613 neural pathways. Various aspects of attention and its influence on perception have been  
614 related to a number of anatomical networks (Shipp, 2004). To date, a dorsal fronto-parietal  
615 network, which entails the intra-parietal sulcus (IPS) in posterior parietal cortex, a portion of  
616 the precentral supplemental motor area, the so-called frontal eye fields (FEF) and early  
617 sensory areas, such as visual cortex has been described most comprehensively (Corbetta and  
618 Shulman, 2002). This cortical network has been implicated in the control of attention  
619 (Corbetta et al., 1998) and was likely involved in deploying the resources necessary to  
620 perform in our behavioral task.

621

622 On the other side, auditory influences on visual processing could have been conveyed by  
623 two candidate routes that have been suggested as a results of earlier invasive  
624 electrophysiological and anatomical studies in the animal brain: (1) feed-forward  
625 projections between thalamus and early sensory cortices (Cappe et al., 2009), (2) lateral  
626 projections between early sensory cortices (Falchier et al., 2002). From our data alone, we  
627 cannot say which pathway was critical in the investigated situation. Both neural pathways  
628 however are anatomically distinct from the fronto-parietal attention network (as described  
629 above) and are thus consistent with our results.

630

631 It should be mentioned that our data analyses and interpretation of results depend on the  
632 implicit assumption that attention and synchrony effects follow similar time courses and,  
633 once established, remain constant through the course of each trial. At least for, spatial  
634 attention we know that gain effects reach asymptote after ~500 ms and keep level for

635 several seconds (Müller et al., 1998b). A time course for synchrony-related gain instead has  
636 not been established yet. This uncertainty notwithstanding, we restricted our analyses to a  
637 period starting 500 ms after stimulus onset. We were confident that this time frame would  
638 allow for enough audio-visual coincidence to be detected to establish synchrony. The  
639 comparison of temporal profiles of attention- and synchrony related gain remains an  
640 interesting subject for future studies, nevertheless.

641

642 As a final remark, Talsma et al. (2010) suggested that bottom-up multisensory integration  
643 benefits a given stimulus the most when competition within one sensory modality is high,  
644 e.g. when the visual field is cluttered. Our situation, with one stimulus presented to each  
645 hemifield, promoted only minimal competition. Inter-hemispheric competition is introduced  
646 relatively late in the visual processing hierarchy (Schwartz et al., 2007). Moreover,  
647 attentional resources seem to split more readily between than within visual hemifields  
648 (Franconeri et al., 2012; Störmer et al., 2013; Walter et al., 2015). It would thus be  
649 interesting to test how synchrony-related gain effects vary with the amount of competition  
650 by placing more than one stimulus within visual hemifields.

651

#### 652 **4.4. Conclusion**

653 We investigated the concurrent effects of spatial attention and audio-visual synchrony on  
654 early cortical visual stimulus processing. Our paradigm allowed us to test both influences in  
655 isolation as well as their combined effects. We found that attention-related and synchrony-  
656 related facilitation add up when an audio-visual synchronous stimulus is attended. Further,  
657 attention facilitated pulse- and flicker-driven neural responses while synchrony only  
658 targeted pulse-driven responses, i.e. those coding for stimulus dynamics that were relevant  
659 for multisensory integration. Consequentially, the present results favor an account in which  
660 goal-directed sustained spatial attention and stimulus-driven audio-visual synchrony convey

661 their influences independently via different neural processes and possibly along different  
662 neural pathways. At least for situations similar to the one studied here, this finding implies  
663 that facilitation through synchrony cannot simply be modelled as a sustained attraction of  
664 spatial attention.

665

#### 666 **Acknowledgments**

667 Work was supported by the Deutsche Forschungsgemeinschaft (grant no. MU972/21-1).

668 Data presented here were recorded at the Institut für Psychologie, Universität Leipzig. The  
669 authors appreciate the assistance of Renate Zahn in data collection. Experimental  
670 stimulation was realized using Cogent Graphics developed by John Romaya at the Laboratory  
671 of Neurobiology at the Wellcome Department of Imaging Neuroscience, University College  
672 London.

673

674 **Conflict of interest:** The authors declare that they have no conflict of interest.

675

676

677 **References**

- 678 Alsius, A., Navarra, J., Campbell, R., Soto-Faraco, S., 2005. Audiovisual Integration of Speech  
679       Falters under High Attention Demands 15, 839–843. doi:10.1016/j.cub.2005.03.046
- 680 Alvarez, G.A., Oliva, A., 2009. Spatial ensemble statistics are efficient codes that can be  
681       represented with reduced attention. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7345–7350.  
682       doi:10.1073/pnas.0808981106
- 683 Andersen, S.K., Hillyard, S.A., Müller, M.M., 2008. Attention Facilitates Multiple Stimulus  
684       Features in Parallel in Human Visual Cortex 18, 1006–1009.  
685       doi:10.1016/j.cub.2008.06.030
- 686 Bundesen, C., Vangkilde, S., Petersen, A., 2015. Recent developments in a computational  
687       theory of visual attention (TVA). *Vision Research* 116, 210–218.  
688       doi:10.1016/j.visres.2014.11.005
- 689 Busse, L., Roberts, K.C., Crist, R.E., Weissman, D.H., Woldorff, M.G., 2005. The spread of  
690       attention across modalities and space in a multisensory object. *Proceedings of the*  
691       *National Academy of Sciences* 102, 18751–18756. doi:10.1073/pnas.0507704102
- 692 Cappe, C., Rouiller, E.M., Barone, P., 2009. Multisensory anatomical pathways. *Hearing*  
693       *Research* 258, 28–36. doi:10.1016/j.heares.2009.04.017
- 694 Chennu, S., Craston, P., Wyble, B., Bowman, H., 2009. Attention Increases the Temporal  
695       Precision of Conscious Perception: Verifying the Neural-ST2 Model. *PLOS Computational*  
696       *Biology* 5, e1000576. doi:10.1371/journal.pcbi.1000576
- 697 Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT Press.
- 698 Colonius, H., Diederich, A., 2012. Focused attention vs. crossmodal signals paradigm:  
699       deriving predictions from the time-window-of-integration model. *Frontiers in*  
700       *Integrative Neuroscience* 6. doi:10.3389/fnint.2012.00062
- 701 Corbetta, M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Ollinger, J.M., Drury, H.A., Linenweber,  
702       M.R., Petersen, S.E., Raichle, M.E., Van Essen, D.C., Shulman, G.L., 1998. A common  
703       network of functional areas for attention and eye movements. *Neuron* 21, 761–773.  
704       doi:10.1016/S0896-6273(00)80593-0
- 705 Corbetta, M., Shulman, G.L., 2002. CONTROL OF GOAL-DIRECTED AND STIMULUS-DRIVEN  
706       ATTENTION IN THE BRAIN. *Nature Reviews Neuroscience* 3, 215–229.  
707       doi:10.1038/nrn755
- 708 Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial  
709       EEG dynamics including independent component analysis. *Journal of Neuroscience*  
710       *Methods* 134, 9–21. doi:10.1016/j.jneumeth.2003.10.009
- 711 Ding, J., 2005. Attentional Modulation of SSVEP Power Depends on the Network Tagged by  
712       the Flicker Frequency. *Cerebral Cortex* 16, 1016–1029. doi:10.1093/cercor/bhj044
- 713 Donohue, S.E., Green, J.J., Woldorff, M.G., 2015. The effects of attention on the temporal  
714       integration of multisensory stimuli. *Frontiers in Integrative Neuroscience* 9.  
715       doi:10.3389/fnint.2015.00032
- 716 Fairhall, S.L., Macaluso, E., 2009. Spatial attention can modulate audiovisual integration at  
717       multiple cortical and subcortical sites. *European Journal of Neuroscience* 29, 1247–  
718       1257. doi:10.1111/j.1460-9568.2009.06688.x
- 719 Falchier, A., Clavagnier, S., Barone, P., Kennedy, H., 2002. Anatomical Evidence of  
720       Multimodal Integration in Primate Striate Cortex. *Journal of Neuroscience* 22, 5749–  
721       5759. doi:10.1002/(SICI)1096-9861(19981026)400:3<417::AID-CNE10>3.0.CO;2-O
- 722 Franconeri, S.L., Scimeca, J.M., Roth, J.C., Helseth, S.A., Kahn, L.E., 2012. Flexible visual  
723       processing of spatial relationships. *Cognition* 122, 210–227.
- 724 Fujisaki, W., Nishida, S., 2005. Temporal frequency characteristics of synchrony–asynchrony  
725       discrimination of audio-visual signals. *Exp Brain Res* 166, 455–464. doi:10.1007/s00221-  
726       005-2385-8

- 727 Giani, A.S., Ortiz, E., Belardinelli, P., Kleiner, M., Preissl, H., Noppeney, U., 2012. Steady-state  
728 responses in MEG demonstrate information integration within but not across the  
729 auditory and visual senses. *NeuroImage* 60, 1478–1489.  
730 doi:10.1016/j.neuroimage.2012.01.114
- 731 Greenhouse, S.W., Geisser, S., 1959. On methods in the analysis of profile data.  
732 *Psychometrika* 24, 95–112. doi:10.1007/BF02289823
- 733 Herrmann, B., Schlichting, N., Obleser, J., 2014. Dynamic Range Adaptation to Spectral  
734 Stimulus Statistics in Human Auditory Cortex. *Journal of Neuroscience* 34, 327–331.  
735 doi:10.1523/JNEUROSCI.3974-13.2014
- 736 Hertz, U., Amedi, A., 2010. Disentangling unisensory and multisensory components in  
737 audiovisual integration using a novel multifrequency fMRI spectral analysis. *NeuroImage*  
738 52, 617–632. doi:10.1016/j.neuroimage.2010.04.186
- 739 Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scandinavian journal*  
740 *of statistics*. doi:10.2307/4615733
- 741 Jenkins, J., Rhone, A.E., Idsardi, W.J., Simon, J.Z., Poeppel, D., 2011. The Elicitation of  
742 Audiovisual Steady-State Responses: Multi-Sensory Signal Congruity and Phase Effects.  
743 *Brain Topography* 24, 134–148. doi:10.1007/s10548-011-0174-1
- 744 Kashiwase, Y., Matsumiya, K., Kuriki, I., Shioiri, S., 2012. Time Courses of Attentional  
745 Modulation in Neural Amplification and Synchronization Measured with Steady-state  
746 Visual-evoked Potentials. *Journal of Cognitive Neuroscience* 24, 1779–1793.  
747 doi:10.1162/jocn\_a\_00212
- 748 Kayser, C., Logothetis, N.K., Panzeri, S., 2010. Visual enhancement of the information  
749 representation in auditory cortex. *Curr. Biol.* 20, 19–24. doi:10.1016/j.cub.2009.10.068
- 750 Keil, J., Mueller, N., Ihssen, N., Weisz, N., 2012. On the Variability of the McGurk Effect:  
751 Audiovisual Integration Depends on Prestimulus Brain States. *Cerebral Cortex* 22, 221–  
752 231. doi:10.1093/cercor/bhr125
- 753 Keitel, C., Andersen, S.K., Quigley, C., Müller, M.M., 2013. Independent Effects of Attentional  
754 Gain Control and Competitive Interactions on Visual Stimulus Processing. *Cerebral*  
755 *Cortex* 23, 940–946. doi:10.1093/cercor/bhs084
- 756 Keitel, C., Müller, M.M., 2015. Audio-visual synchrony and feature-selective attention co-  
757 amplify early visual processing. *Exp Brain Res* 234, 1221–1231. doi:10.1007/s00221-015-  
758 4392-8
- 759 Kim, Y.J., Grabowecy, M., Paller, K.A., Muthu, K., Suzuki, S., 2007. Attention induces  
760 synchronization-based response gain in steady-state visual evoked potentials. *Nature*  
761 *Neuroscience* 10, 117–125. doi:10.1038/nn1821
- 762 Kim, Y.J., Grabowecy, M., Paller, K.A., Suzuki, S., 2011. Differential Roles of Frequency-  
763 following and Frequency-doubling Visual Responses Revealed by Evoked Neural  
764 Harmonics. *Journal of Cognitive Neuroscience* 23, 1875–1886.  
765 doi:10.1162/jocn.2010.21536
- 766 Krause, H., Schneider, T.R., Engel, A.K., Senkowski, D., 2012. Capture of visual attention  
767 interferes with multisensory speech processing. *Frontiers in Integrative Neuroscience* 6.  
768 doi:10.3389/fnint.2012.00067
- 769 Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., Schroeder, C.E., 2008. Entrainment of  
770 Neuronal Oscillations as a Mechanism of Attentional Selection. *Science* 320, 110–113.  
771 doi:10.1126/science.1154735
- 772 Lakatos, P., O'Connell, M.N., Barczak, A., Mills, A., Javitt, D.C., Schroeder, C.E., 2009. The  
773 Leading Sense: Supramodal Control of Neurophysiological Context by Attention. *Neuron*  
774 64, 419–430. doi:10.1016/j.neuron.2009.10.014
- 775 Lee, S., 1999. Visual Form Created Solely from Temporal Structure. *Science* 284, 1165–1168.  
776 doi:10.1126/science.284.5417.1165
- 777 Leek, M.R., 2001. Adaptive procedures in psychophysical research. *Perception &*



- 778 Psychophysics 63, 1279–1292. doi:10.3758/BF03194543
- 779 Morey, R.D., Rouder, J.N., Jamil, T., (2015) BayesFactor: Computation of Bayes Factors for  
780 Common Designs. R package. URL <http://bayesfactorppl.r-forge.r-project.org/>
- 781 Morgan, S.T., Hansen, J.C., Hillyard, S.A., 1996. Selective attention to stimulus location  
782 modulates the steady-state visual evoked potential. *Proceedings of the National*  
783 *Academy of Sciences* 93, 4770–4774. doi:10.1073/pnas.93.10.4770
- 784 Müller, M.M., Hubner, R., 2002. Can the Spotlight of Attention be Shaped Like A Doughnut?  
785 *Psychological Science* 13, 119–124. doi:10.1111/j.0956-7976.2002.t01-1-x
- 786 Müller, M.M., Picton, T.W., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W.A., Hillyard, S.A.,  
787 1998a. Effects of spatial selective attention on the steady-state visual evoked potential  
788 in the 20–28 Hz range. *Cognitive Brain Research* 6, 249–261. doi:10.1016/S0926-  
789 6410(97)00036-0
- 790 Müller, M.M., Teder-Sälejärvi, W., Hillyard, S.A., 1998b. The time course of cortical  
791 facilitation during cued shifts of spatial attention. *Nature Neuroscience* 1, 631–634.  
792 doi:10.1038/2865
- 793 Nolan, H., Whelan, R., Reilly, R.B., 2010. FASTER: Fully Automated Statistical Thresholding for  
794 EEG artifact Rejection. *Journal of Neuroscience Methods* 192, 152–162.  
795 doi:10.1016/j.jneumeth.2010.07.015
- 796 Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottureau, B.R., Rossion, B., 2015. The steady-  
797 state visual evoked potential in vision research: A review. *Journal of Vision* 15, 4–4.  
798 doi:10.1167/15.6.4
- 799 Nozaradan, S., Peretz, I., Mouraux, A., 2012. Steady-state evoked potentials as an index of  
800 multisensory temporal binding. *NeuroImage* 60, 21–28.  
801 doi:10.1016/j.neuroimage.2011.11.065
- 802 Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: Open source software  
803 for advanced analysis of MEG, EEG, and invasive electrophysiological data.  
804 *Computational Intelligence and Neuroscience* 2011, 156869–9.  
805 doi:10.1155/2011/156869
- 806 Parise, C.V., 2012. Signal compatibility as a modulatory factor for audiovisual multisensory  
807 integration.
- 808 Parise, C.V., Ernst, M.O., 2016. Correlation detection as a general mechanism for  
809 multisensory integration. *Nature Communications* 7, 11543. doi:10.1038/ncomms11543
- 810 Pastor, M.A., Valencia, M., Artieda, J., Alegre, M., Masdeu, J.C., 2007. Topography of cortical  
811 activation differs for fundamental and harmonic frequencies of the steady-state visual-  
812 evoked responses. An EEG and PET H215O study. *Cerebral Cortex* 17, 1899–1905.  
813 doi:10.1093/cercor/bhl098
- 814 Porcu, E., Keitel, C., Müller, M.M., 2014. Visual, auditory and tactile stimuli compete for early  
815 sensory processing capacities within but not between senses. *NeuroImage* 97, 224–235.  
816 doi:10.1016/j.neuroimage.2014.04.024
- 817 Porcu, E., Keitel, C., Müller, M.M., 2013. Concurrent visual and tactile steady-state evoked  
818 potentials index allocation of inter-modal attention: A frequency-tagging study.  
819 *Neuroscience Letters* 556, 113–117. doi:10.1016/j.neulet.2013.09.068
- 820 Posner, M.I., Snyder, C.R., Davidson, B.J., 1980. Attention and the detection of signals.  
821 *Journal of Experimental Psychology: General* 109, 160–174. doi:10.1037/0096-  
822 3445.109.2.160
- 823 Quigley, C., Müller, M.M., 2014. Feature-Selective Attention in Healthy Old Age: A Selective  
824 Decline in Selective Attention? *Journal of Neuroscience* 34, 2471–2476.  
825 doi:10.1523/JNEUROSCI.2718-13.2014
- 826 R Core Team (2013). R: A language and environment for statistical computing. R Foundation  
827 for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL [http://www.R-](http://www.R-project.org/)  
828 [project.org/](http://www.R-project.org/).



- 829 Regan, D., 1989. Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic  
830 Fields in Science and Medicine. (1989).
- 831 Reynolds, J.H., Heeger, D.J., 2009. The Normalization Model of Attention. *Neuron* 61, 168–  
832 185. doi:10.1016/j.neuron.2009.01.002
- 833 Rouder, J.N., Morey, R.D., Speckman, P.L., Province, J.M., 2012. Default Bayes factors for  
834 ANOVA designs. *Journal of Mathematical Psychology* 56, 356–374.
- 835 Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., Iverson, G., 2009. Bayesian t tests for  
836 accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review* 16, 225–  
837 237. doi:10.3758/PBR.16.2.225
- 838 Ruhнау, P., Keitel, C., Lithari, C., Weisz, N., Neuling, T., 2016. Flicker-Driven Responses in  
839 Visual Cortex Change during Matched-Frequency Transcranial Alternating Current  
840 Stimulation. *Frontiers in Human Neuroscience* 10, 440. doi:10.3389/fnhum.2016.00184
- 841 Schall, S., Quigley, C., Onat, S., König, P., 2009. Visual stimulus locking of EEG is modulated by  
842 temporal congruency of auditory stimuli. *Exp Brain Res* 198, 137–151.  
843 doi:10.1007/s00221-009-1867-5
- 844 Schönbrodt, F.D., Wagenmakers, E.J., 2015. Sequential hypothesis testing with Bayes factors:  
845 Efficiently testing mean differences. *Psychological ...*
- 846 Schwartz, O., Hsu, A., Dayan, P., 2007. Space and time in visual context. *Nature Reviews*  
847 *Neuroscience* 8, 522–535. doi:10.1038/nrn2155
- 848 Shipp, S., 2004. The brain circuitry of attention. *Trends in Cognitive Sciences* 8, 223–230.  
849 doi:10.1016/j.tics.2004.03.004
- 850 Störmer, V., Cavanagh, P., Alvarez, G., 2013. The profile of multifocal attention: surround-  
851 suppression between and within hemifields. *Journal of Vision* 13, 1283–1283.  
852 doi:10.1167/13.9.1283
- 853 Störmer, V.S., Alvarez, G.A., Cavanagh, P., 2014. Within-Hemifield Competition in Early Visual  
854 Areas Limits the Ability to Track Multiple Objects with Attention. *Journal of*  
855 *Neuroscience* 34, 11526–11533. doi:10.1523/JNEUROSCI.0980-14.2014
- 856 Śmigasiewicz, K., Asanowicz, D., Westphal, N., Verleger, R., 2014. Bias for the Left Visual  
857 Field in Rapid Serial Visual Presentation: Effects of Additional Salient Cues Suggest a  
858 Critical Role of Attention. *Journal of Cognitive Neuroscience* 27, 266–279.  
859 doi:10.1162/jocn\_a\_00714
- 860 Talsma, D., Senkowski, D., Soto-Faraco, S., Woldorff, M.G., 2010. The multifaceted interplay  
861 between attention and multisensory integration. *Trends in Cognitive Sciences* 14, 400–  
862 410. doi:10.1016/j.tics.2010.06.008
- 863 Talsma, D., Senkowski, D., Woldorff, M.G., 2009. Intermodal attention affects the processing  
864 of the temporal alignment of audiovisual stimuli. *Exp Brain Res* 198, 313–328.  
865 doi:10.1007/s00221-009-1858-6
- 866 Treisman, A.M., Gelade, G., 1980. A feature-integration theory of attention. *Cognitive*  
867 *Psychology* 12, 97–136. doi:10.1016/0010-0285(80)90005-5
- 868 van Atteveldt, N., Murray, M.M., Thut, G., Schroeder, C.E., 2014. Multisensory Integration:  
869 Flexible Use of General Operations. *Neuron* 81, 1240–1253.  
870 doi:10.1016/j.neuron.2014.02.044
- 871 Van der Burg, E., Olivers, C.N.L., Bronkhorst, A.W., Theeuwes, J., 2008. Pip and pop:  
872 Nonspatial auditory signals improve spatial visual search. *Journal of Experimental*  
873 *Psychology: Human Perception and Performance* 34, 1053–1065. doi:10.1037/0096-  
874 1523.34.5.1053
- 875 Verleger, R., Śmigasiewicz, K., Möller, F., 2011. Mechanisms underlying the left visual-field  
876 advantage in the dual stream RSVP task: Evidence from N2pc, P3, and distractor-evoked  
877 VEPs. *Psychophysiology* 48, 1096–1106. doi:10.1111/j.1469-8986.2011.01176.x
- 878 Walter, S., Keitel, C., Müller, M.M., 2015. Sustained Splits of Attention within versus across  
879 Visual Hemifields Produce Distinct Spatial Gain Profiles. *Journal of Cognitive*

- 880 Neuroscience 28, 111–124. doi:10.1162/jocn\_a\_00883
- 881 Walter, S., Quigley, C., Andersen, S.K., Mueller, M.M., 2012. Effects of overt and covert  
882 attention on the steady-state visual evoked potential. *Neuroscience Letters* 519, 37–41.  
883 doi:10.1016/j.neulet.2012.05.011
- 884 Werner, S., Noppeney, U., 2011. The contributions of transient and sustained response  
885 codes to audiovisual integration. *Cereb. Cortex* 21, 920–931.  
886 doi:10.1093/cercor/bhq161
- 887 Wolfe, J.M., 1994. Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin*  
888 *& Review* 1, 202–238. doi:10.3758/BF03200774
- 889 Wolfe, J.M., Cave, K.R., Franzel, S.L., 1989. Guided search: An alternative to the feature  
890 integration model for visual search. *Journal of Experimental Psychology: Human*  
891 *Perception and Performance* 15, 419–433. doi:10.1037/0096-1523.15.3.419
- 892 Zhang, D., Hong, B., Gao, X., Gao, S., Röder, B., 2010. Exploring steady-state visual evoked  
893 potentials as an index for intermodal and crossmodal spatial attention.  
894 *Psychophysiology* 48, 665–675. doi:10.1111/j.1469-8986.2010.01132.x
- 895

896 **Figure captions**

897

898 **Figure 1** Stimulation details. (A) On-screen stimulus display comprising central fixation rings  
 899 and one Gabor patch per lower left and right visual hemifield. All items not to scale.  
 900 Participants received auditory stimulation via headphones. (B) Schematic trial time course.  
 901 An instructive position cue allocates attention to the left or right stimulus. Subsequent  
 902 ongoing Gabor-patch and tone stimulation are represented by grey sinusoids. (C) A common  
 903 frequency modulation (FM; solid black line) of auditory tone pitch and the spatial frequency  
 904 of one of the two Gabor patches produces a synchronous pulsing audio-visual percept.  
 905 Concurrently, the spatial frequency of the other Gabor patch modulates at a slightly  
 906 different frequency (dashed grey line), thus rendering it asynchronous to the tone.  
 907 (D) Frame-by-frame visual stimulation for the right Gabor patch. The illustration shows the  
 908 first 27 frames of each trial. Note the emphasis on the on-off cycles leading to a 17-Hz flicker  
 909 along the horizontal axis (black boxes = off-frames) and one full cycle of the spatial  
 910 frequency modulation leading to a 3.14-Hz ‘pulsation’ along the vertical axis.

911

912 **Figure 2** Stimulus-driven steady-state responses (SSRs) – spectra and scalp maps. (A) SSR  
 913 power extracted from spectral decomposition of trial-averaged EEG waveforms, thus  
 914 “stimulus-evoked”. Scalp maps show topographical distributions of power for the pulse-  
 915 frequency following (*pulse 1f*), pulse-frequency doubling (*pulse 2f*) and flicker-frequency  
 916 following (*flicker 1f*) SSR components driven by left and right stimuli respectively. White dots  
 917 in left-most scalp map highlight the uniform sensor cluster used in all data analyses. Spectra  
 918 below depict condition-averaged individual power spectra (grey lines) and, superimposed in  
 919 black, the grand-average spectrum. Arrows indicate peaks that correspond to the respective  
 920 driving frequencies (in Hz). (B) Same as (A) but for SSR inter-trial phase consistency (ITC)  
 921 measured in arbitrary units (au). (C) Power spectra based on averaged spectral  
 922 decompositions of single trials for comparison. Note that this approach emphasizes spectral  
 923 characteristics of the ongoing EEG, such as the alpha rhythm (see peaks around 10 Hz,  
 924 denoted  $\alpha$ ), over SSRs given our stimulation.

925

926 **Figure 3** SSRs by condition. (A) Condition-resolved grand-average power (dB) spectra. Top  
 927 panel: Spectra split for Attend Left (dark graph) and Attend Right (light graph) conditions.  
 928 Bottom panel: Spectra split for conditions in which the tone pulsed in synchrony with the left  
 929 (dark) or right (light) Gabor patch. Shaded areas represent standard error of the mean  
 930 (SEM). Arrows pointing to peaks indicate the spatial position of the corresponding driving  
 931 stimulus (L = left, R = right). (B) Same as in (A) but for SSR inter-trial phase coherence (ITC)  
 932 measured in arbitrary units (au). (C) Zoom-in on power at SSR component frequencies. For  
 933 each frequency, box plots showcase inter-individual power distributions. Boxes depict  
 934 interquartile ranges with medians superimposed as strong horizontal lines. Grey dots signify  
 935 outliers. A common color code applies (also see color key): Hot colors = corresponding visual  
 936 stimulus attended; Monochrome = visual stimulus unattended; Light colors = visual stimulus  
 937 in sync with tone; Dark colors = visual stimulus and tone asynchronous. (D) Same as in C but  
 938 for SSR inter-trial coherence.

939

940 **Figure 4** Quantifying and comparing attention- and synchrony related gain modulation. (A)  
 941 SSR power (in dB) for all three SSR components of interest (*pulse 1f*, *pulse 2f* and *flicker 1f*)  
 942 separated by whether the driving visual stimulus was attended (orange) or unattended (red).  
 943 Box plots display inter-individual power distributions. Boxes depict respective interquartile  
 944 ranges with medians superimposed as strong horizontal lines. (B) Same as in (A) but for SSR  
 945 inter-trial phase coherence (ITC) measured in arbitrary units (au). (C) SSR power (in dB) for  
 946 *pulse 1f*, *pulse 2f* and *flicker 1f* components separated by whether the driving visual stimulus

947 pulsed in sync with the tone (light grey) or asynchronous (dark grey). (D) Same as in (C) but  
948 for SSR inter-trial phase coherence (ITC) measured in arbitrary units (au). (E) Boxes indicate  
949 SSR power modulation (in au) by attention (brown) and synchrony (blue) for *pulse 1f*,  
950 *pulse 2f* and *flicker 1f* components of interest. (F) Same as in (C) but for modulation of SSR  
951 inter-trial phase coherence (in au). Grey dots in plots signify outlier values. Asterisks close to  
952 medians in E & F demarcate statistically significant deviations from zero, i.e. systemic gain  
953 modulations (two-tailed t-tests,  $P < .05$ , Holm-Bonferroni corrected for multiple  
954 comparisons).  
955

ACCEPTED MANUSCRIPT

**Table 1** Average behavioral performance in the visual fading detection task (N = 12).

Attended Stimulus		Left		Right	
		S+	S-	S+	S-
Proportion	<i>M</i>	85.6 %	84.2 %	76.4 %	76.8 %
correct (%)	$\pm$ SEM	2.2 %	2.0 %	2.4 %	2.7 %
Reaction	<i>M</i>	674	662	667	662
time (ms)	$\pm$ SEM	14	16	16	13

M = mean; SEM = standard error of the mean; S+ = synchronous; S- = asynchronous

